

BAB 7 PENUTUP

Bab ini menjelaskan kesimpulan dari penelitian ini dan memberikan saran untuk penelitian selanjutnya.

7.1 Kesimpulan

Berdasarkan hasil pengujian dan analisis dari penelitian Analisis Sentimen Kurikulum 2013 pada Twitter menggunakan *Ensemble Feature* dan metode *K-Nearest Neighbor* (K-NN) maka diperoleh kesimpulan sebagai berikut:

1. Metode *K-Nearest Neighbor* (K-NN) dengan fitur *ensemble* secara lengkap dapat diimplementasikan untuk mengklasifikasikan *tweet* positif dan negatif dengan akurasi diatas 80%. Adapun alur dalam mengklasifikasikan *tweet* tersebut adalah sebagai berikut:
 - a. Memasukkan *datasets* yaitu data *tweet* latih dan data *tweet* uji ke dalam *database*.
 - b. Melakukan proses pembakuan kata-kata tertentu pada *tweet* dan *preprocessing* data latih dan data uji yang meliputi proses *tokenization*, *filtering/stopword*, *stemming*.
 - c. Melakukan proses pengambilan nilai fitur pada data latih dan data uji, pengambilan nilai fitur meliputi nilai dari fitur *Bag of Words* (BoW) dan fitur *ensemble* (*twitter specific features*, *textual features*, *PoS features*, *lexicon based features*).
 - d. Melakukan proses testing dengan data uji, yaitu dengan menghitung *cosine similarity* atau kemiripan data uji terhadap data latih yang kemudian akan diklasifikasikan menggunakan metode *K-Nearest Neighbor* (K-NN) dengan mengambil nilai *k* tertentu. Dari hasil testing tersebut maka didapatkan hasil klasifikasi *tweet* uji.
 - e. Menghitung nilai akurasi yang didapatkan setelah dilakukannya proses *testing*.
2. Nilai *k* sangat berpengaruh penting dalam akurasi metode *K-Nearest Neighbor* (K-NN), nilai *k* terbaik yang didapatkan setelah dilakukan pengujian adalah 5. Nilai *k* tersebut didapatkan dengan menguji nilai *k* secara berganti-ganti dan dilakukan pengambilan nilai *k* dengan akurasi tertinggi yaitu, sebesar 96%. Nilai *k* yang terlalu sedikit menyebabkan akurasi yang didapatkan belum mencapai titik maksimal sebaliknya, nilai *k* yang terlalu banyak akan menyebabkan akurasi menurun.
3. Variasi atau kombinasi fitur memiliki pengaruh penting yang signifikan dalam meningkatkan akurasi dari metode *K-Nearest Neighbor* (K-NN), penggabungan fitur *Bag of Words* (BoW) dan *ensemble* (*twitter specific features*, *textual features*, *PoS features*, *lexicon based features*) dapat meningkatkan akurasi metode *K-Nearest Neighbor* (K-NN) jika

dibandingkan dengan menggunakan fitur secara independen saja. Penggabungan fitur ini dapat melengkapi kelemahan masing-masing fitur, akurasi yang didapatkan dengan menggabungkan kedua fitur tersebut mencapai 96%.

7.2 Saran

Berdasarkan penelitian yang telah dilakukan, didapatkan beberapa saran yang dapat dikembangkan untuk penelitian selanjutnya, berikut saran-saran yang didapatkan:

1. Melakukan penambahan data pada kamus *emoticon*, kamus kata baku dan tidak baku, data kamus *intensifier word* (kata penguat), kata dasar dan kamus *Parts of Speech (PoS)* berbahasa Indonesia.
2. Melakukan penelitian dengan menggunakan metode klasifikasi yang berbeda.
3. Melakukan penambahan jumlah data *tweet* latih dan data *tweet* uji dengan harapan akan dapat meningkatkan akurasi dalam analisis sentimen.